



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2007

How up-to-date are online tourism communities? An empirical evaluation of commercial and non-commercial information quality

Prestipino, Marco ; Aschoff, Felix-Robinson ; Schwabe, Gerhard

Abstract: While social and economic aspects of online communities have been investigated broadly, the information exchanged has seldom been the subject of study. The article follows recent work on using an information systems metaphor for online communities: users specify queries and receive information from the online community members. In order to justify this metaphor, information needs to be at least up to classic information products. In this paper we present a framework for the evaluation of timeliness in online communities. An empirical study is presented which compares aspects of timeliness, namely up-to-dateness, for a wiki community and a printed guidebook. Results show that the community is at least as up-to-date as the printed guidebook. While further research is needed, results indicate that online communities can be used as information systems with reasonable information quality values.

DOI: <https://doi.org/10.1109/HICSS.2007.251>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-61380>

Conference or Workshop Item

Published Version

Originally published at:

Prestipino, Marco; Aschoff, Felix-Robinson; Schwabe, Gerhard (2007). How up-to-date are online tourism communities? An empirical evaluation of commercial and non-commercial information quality. In: 40. HICCS Hawaii International Conference on System Sciences, Hawaii, 3 January 2007 - 6 January 2007, IEEE.

DOI: <https://doi.org/10.1109/HICSS.2007.251>

How up-to-date are Online Tourism Communities?

An Empirical Evaluation of Commercial and Non-commercial Information Quality

Marco Prestipino, Felix-Robinson Aschoff and Gerhard Schwabe

*Department of Informatics
University of Zurich, Switzerland
{prestipino, aschoff, schwabe}@ifi.unizh.ch*

Abstract

While social and economic aspects of online communities have been investigated broadly, the information exchanged has seldom been the subject of study. The article follows recent work on using an information systems metaphor for online communities: users specify queries and receive information from the online community members. In order to justify this metaphor, information needs to be at least up to classic information products. In this paper we present a framework for the evaluation of timeliness in online communities. An empirical study is presented which compares aspects of timeliness, namely up-to-dateness, for a wiki community and a printed guidebook. Results show that the community is at least as up-to-date as the printed guidebook. While further research is needed, results indicate that online communities can be used as information systems with reasonable information quality values.

1. Introduction

Online Communities (or virtual communities) are now a broadly discussed and researched phenomenon ([14], [17], [8], [23]). But their role in knowledge management has seldom been discussed outside of corporate online communities of practice, although a myriad of much larger and active communities exist open on the internet. The long history of discussion spaces on the internet and also new developments, such as the tremendous success of the community-created encyclopedia Wikipedia reflect the potential of online communities to serve the information needs of its users much like a traditional information product but with a “social” retrieval mechanism. As argued in [15], online communities provide not only places for socializing and giving social support to their members,

but also create and distribute valuable information, and could be seen as information systems in the sense that users may formulate an information need and retrieve information. Hence it must be possible to assess the performance of an online community in terms of information quality. To justify the concept of online communities as information systems, information in online communities and access to it must be at least as good as when using classic information products. Many features of online communities seemingly support this claim: online communities provide a “natural language”-interface to information, there is no need to transform an information need into a formalized query language, thereby loosing details of the information need, or using fixed navigational structures, e.g. indices. The community acts proactively and may provide information the asking person did not think of or deem necessary or failed to formulate in his query. As in an online community the distinction between writers and consumers is blurred, a large number of authors may create content, distributing the work load.

Empirically, we observe the existence of large and heavily used online communities with specific thematic foci, supporting the claim that they do not only serve socializing needs. As an example, the newsgroup “rec.travel.europe”, existing since 1994, had 8192 contributors writing 97’289 messages in 2003¹.

Consequently, the quality of information is an important factor when selecting the right online community for information needs, and may determine a community’s success. While little research has been published on information quality in large open online communities, Neus [12] notes that in small scale online

¹<http://netscan.research.microsoft.com/Text/?Timespan=Y&SearchDate=1/1/2003&SearchFor=rec.travel.europe&NGID=2060&Metric=Messages>, as of 06/16/2006

communities, the issue of information quality is often ignored, or the community is overly controlled and supported by poor tools not designed for social interaction. Few tools (cf. [18]) or frameworks have been published specifically for the assessment of information quality in online communities on the web.

In this article, we explore the issue of the timeliness dimension in online communities, discuss differences to information products and present empirical data on the up-to-dateness found in an online community for travelers and a guidebook. We chose the domain of independent travel because it is an information intensive activity, with travelers frequently needing to take decisions and thus needing information. These information needs are also very individual and complex, as individual preferences along with the circumstances of the requester need to be accounted for. Stockdale & Borovicka [22] analyze Lonely Planet's online community and conclude that "Visitors to the site can find a wealth of information on a range of travel related topics [...] Lonely Planet realizes several benefits from the community. They have a constant flow of information, albeit unstructured and often effusive, and feedback." However, these findings are merely based on observation by the authors and not backed up with empirical data about community activity and information quality.

2. Related Work

The notion of information quality has been widely discussed in literature. Several quality attributes, dimensions and frameworks have been proposed. Alter [1] lists relevant quality attributes: types of data, accuracy/precision, age, time horizon, level of summarization, completeness, accessibility, source as well as relevance/value. Senn [21] refers to the attributes of accuracy, form, frequency, breadth, origin, and time horizon. More recent approaches propose frameworks on more abstract levels which specify algebraic relations for different information quality dimensions ([4], [19]). In many studies about information quality, quality attributes are selected based on the researcher's intuition, which allows selecting the criteria relevant to the study's goals [9]. On the other hand, the intuitive approach provides quality attributes for a production perspective, and does not account for the user perspective. Burgess, Fiddian et al. [3] also conclude that information quality is mostly analyzed from an organizational and information producer's perspective, while the user perspective is neglected. Our concept of information quality is based on the notion that "Information is knowledge in action" [10], meaning that information

exists only in a certain context and with regard to a certain need. Thus, information quality is mainly to be evaluated from the user's point of view.

While the different conceptualizations of information quality found in literature can hardly be integrated into one consistent framework, they also lack in many cases concrete procedures on how to empirically assess the quality attributes in practice [7]. Eppler [5] also notes the lack of tools to support an assessment of information quality. Definitions of quality attributes are often short and open for debate.

3. Timeliness as quality dimension

Based on the notion of a knowledgeable traveler we have defined four dimensions of information quality relevant to a traveler before and during a journey [20]: Completeness, Structure, Personalization and Timeliness. *Completeness* refers to the degree to which a medium is able to serve a user's information need, implicitly capturing other criteria as ease of understanding, and serving as an indicator for relevancy (for an empirical evaluation see [16]). *Structure* refers to the presentation and structure of information, which may greatly affect efficiency of information access and learning and is therefore closely related to the criteria accessibility and understandability found in literature. *Personalization* indicates how fitting the information is for a person in her real world context, and how much unnecessary and unsuitable information is returned to her. Finally, *timeliness* refers to whether information is up-to-date and available to the user in an acceptable timeframe. Timeliness is mentioned in most frameworks on information quality (cf. [5]), and auditing timeliness of existing information systems is an open research problem and critical for management of information production [2].

Poor information quality in an online community lessens the community's value and the efficiency of the communication therein. Users may switch to other information sources if the effort does not justify the results, so the user perspective is paramount for information quality in an online community.

2.1 Attributes for the timeliness dimension

In this section, we propose attributes for the concept of timeliness from a user perspective and describe how they relate to information products and online communities.

Up-to-dateness refers to the information either retrieved from the forum by an automated search tool or received as answer to a posted question. Up-to-dateness describes the fit between an object or event in the real world and its description given in the information. Up-to-date information is correct, as information that does not fit the real world is not correct. (As an example, if a phone number of a hotel is given as 0433092, and the hotel really is reached by number 0394442, the information is neither up-to date nor correct). The opposite of up-to-date information is incorrect, obsolete information. Information may also never have been up-to-date, in which case it was already incorrect by the time it was created, e.g. entered into an information system. The age of information in a system is at most an indicator for up-to-dateness, as up-to-dateness is dependent on the frequency of changes or creation of new information in a source. The importance of up-to-dateness from a user perspective is influenced by whether the information is needed to complete a task at hand or for learning without immediate usage of that information.

Speed measures how long it takes until information to a posted question is received. Information should be available when it is needed. In the traditional context of information systems, speed can be determined by evaluating the performance of the algorithms and the hardware used, and the transmission speed of the communication means between the IS and user. In an online community, however, speed is less dependent on technical parameters than on the activity of its members. As information is distributed and created in a communication and cooperation process, speed may be described by two measures: the time span until the first reply to a posted question is received and the time span until the last post is received. These measures need to be determined by observation and may only be given as an average. Speed is not concerned with the completeness or suitability of the information received, but with the provision process. Information returned may be useless or incomplete, but this is partly influenced by the competency of the user to formulate a query in both the IS and the online community. In a qualitative observation, speed may additionally be measured by the time span until the first and last information deemed suitable is returned or the time span until the frequency of answers begins to decline, as indicators for the optimum time to wait for information returned by the community.

Note that a non-interactive usage of an online community, e.g. by searching the archive, is similar to conventional usage of information products, e.g. books or IS.

Time-to-publish describes how timely newly observed information is made accessible for possible recipients. As an example, if a reader writes updated or new information for a printed guidebook to a publisher, it may only be visible in the next edition of the guidebook. This is especially important for urgent, time-dependent information needs, e.g. coordination tasks. In an online community, Time-to-publish is also the time span after which answers to input is visible to other users, i.e. the velocity of transmitting a question or an answer to an information request usually equals Time-to-publish. In an IS this time span would usually equal the time for the speed attribute.

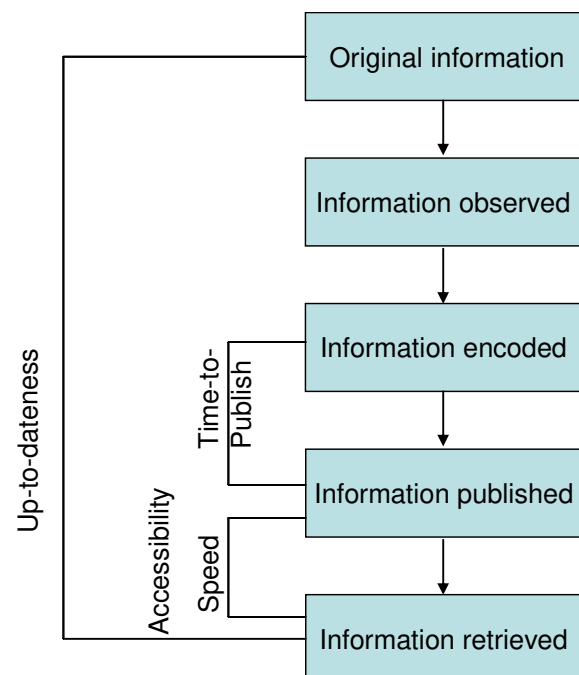


Figure 1. Elements of Timeliness

Accessibility, defined as the „capability of being reached, capable of being seen or used” [5], may also be considered for the timeliness dimension, but introduces issues much more dependent on the individual user and the content. A user with high competency in information retrieval may retrieve information much faster than a novice. Similarly, good structure of content may reduce search time significantly. In an online community a user does not have to use any restricted query language or mechanism, as he poses questions in natural language. Kahn et al. [9] describe accessibility as: „the extent to which information is available, or easily and quickly retrievable”. While no further explanation is given, the extent to which information is quickly retrievable is

certainly related to the timeliness dimension. But as it is difficult to measure in itself, we treat it as a set of factors influencing the speed attribute. Of course a guidebook is easy to be taken along, so its accessibility may be higher in this regard. But with the advent of ever more powerful mobile devices, this may change. Figure 1 depicts how our attributes describe timeliness at different stages of an information creation and distribution process and how speed is influenced by accessibility.

To allow for a meaningful interpretation of timeliness measurements, we compare online communities to a standard benchmark. The standard on the market for travel information is the most popular information product for travellers, the printed guidebook.

2.2 Timeliness in information creation and distribution

This section will describe timeliness in the processes of information creation and distribution for an information product and an online community. The process of creating an information product – a guidebook in our case – involves a sequence of several phases:

- 1) Research and writing phase, in which information is collected and condensed in structured text.
- 2) Printing and deployment phase, in which books are produced and shipped to vendors.
- 3) Selling and usage phase during which time and consumers may obtain the book until it is no longer available and possibly replaced by a newer edition. Usage of a particular edition, of course, continues long after it has been out of print. As the research and writing phase is carried out by a limited number of people, it spans over a non-negligible period of time, meaning that some information collected in the beginning may already be outdated towards the end of the phase. Printing and deployment may take several months.

To cope with the tremendous amount of information, some information may be poorly researched or taken from previous editions without sufficient checking. Typing, translation or communication errors may occur during creation of a manuscript. Hence, even when deployed, a guidebook contains information that is outdated or was incorrect in the first place. Afterwards, the guidebook's up-to-

dateness is subject to continuous degradation. As the information product grows older, more and more information is outdated and thus wrong. Of course, the volatility of information units varies, e.g. historical, geographical or cultural information has a longer lifespan than accommodation or transportation information. The publisher may produce a new edition, and the cycle begins anew. Up-to-dateness is dependent on the time span between updates and the frequency of changes to the information contained in the guidebook in the real world. The obvious solutions would be to

- a) shorten the lifetime of editions to improve up-to-dateness, but this would cut down on revenue per edition (assuming sales would not increase significantly), or publishers would need to justify a higher price. Because of limited resources, a reduced time span may be insufficient for thorough research, lowering information quality.
- b) include only information that changes with a frequency suited for the production cycle, leaving out more volatile information.

In an online community, information is usually immediately available through the system after it has been written; creation and deployment do not delay the usage phase noticeably. Information may be added, commented or changed directly dependent on the technology. Outdated information may thus be updated and corrected at any time, as there are no discrete editions. Content creation in communities is decentralized to an unlimited number of authors, as there is no separation between readers and writers. So the gap between real-world observation and writing (and publishing) may be much shorter than with books. An unlimited number of contributors may participate. While the author of a printed guidebook has to report his findings back to the editorial office where his information undergoes further refinement, a traveler on a journey can immediately update information of the community in any internet café and, thus, make it available to all community members.

Consequently, online communities are capable of mediating short-lived information that is too volatile to be distributed via information products; examples include weather conditions or local events where time between observation and obsolescence is short. On the other hand the externalized information in a community system is not updated in a formalized process, such as editions for a book. There are no guidelines for quality. Thus information units may differ substantially in up-to-dateness. While authors of

a guidebook receive monetary compensation and thus may have an obligation to provide a certain comprehensiveness and quality content, money does not lead to intrinsic motivation for knowledge based work ([6], [13]).

Another major difference is that online communities offer an additional retrieval mechanism: besides using information units already available in the community archives, a user may ask for information in a board. Other community members may then answer his request. The time span between request and answer is determined by the activity of the community, i.e. how many users read requests and discussions and then write answers in a given time span. However, this time span does not necessarily indicate the up-to-dateness of the information, as the person writing the answer may have acquired this information long ago, for example on a trip he once made, without the ability to recheck if it is still correct.

As a community does not release information at defined intervals, its up-to-dateness is not influenced by publishing intervals, but depends on the activity of its members and its motivation and ability to keep the content up-to-date. Information is only partially stored in databases, as the externalized information visible in the system represents only part of the knowledge the online community has through its members. Information is added to the system on request, e.g. following a question in a discussion board or as a reaction to the discovery of incomplete or erroneous information in a wiki. Hence, it is not only the influence of the (technical) delivery mechanism influencing but also the speed of the creation process that determines overall speed. Lonely Planet uses a drastic way of keeping its discussion forum's up-to-dateness high: we found that most posts are deleted after a few weeks.

Table 1 summarizes the differences between the guidebook and the online community. It is interesting to see that while Time-to-publish in online communities is negligibly low, it is not in guidebooks, while it is the opposite for the speed attribute. In a guidebook, information that is available is usually found quickly, influenced by accessibility features like structure. In an online community, it depends on the activity profile, i.e. how many users read and write and how often they visit the online community and how quickly they react. Accessibility has a bundle of factors influencing timeliness attributes. In an online community, natural language, proactivity and feedback are the main factors, while the guidebook can easily be taken along by the traveler and has a proven structure allowing for quick access to elements included in the structure. The structure may even be improved by

manipulating the guidebook, e.g. tearing out or annotating pages.

Table 1. Elements of timeliness in guidebooks and online communities

	<i>Guidebooks, information products</i>	<i>Online community</i>
Time-to-publish	Dependent on resources for creating new editions on a profitable basis, usually several years	Dependent on technical factors, usually negligible
Up-to-dateness	Dependent on resources invested in creation and age of edition	Dependent on community activity
Speed	Dependent on technical factors (network bandwidth, hardware), low if information is available	Dependent on community activity
Accessibility	Structured content, non-interactive, no tailoring of information, portable	Natural language retrieval, tailoring of information, access through standard internet connection

2.3 Hypothesis

The rest of this article will focus on assessing up-to-dateness in an information product and an online community. Up-to-dateness is a central attribute with a high influence on the overall value for the information user; information with a low up-to-dateness may be considered useless in spite of being high in other attributes. While assessment of speed is relatively easy to measure and can be automated in computer mediated communication systems, assessment of up-to-dateness requires a complex design. We are not aware of any other studies of up-to-dateness in online communities.

Our research interest lies in the question whether a non-commercial, non-centrally organized organisation like an online community may be capable of producing

information which is satisfying in terms of up-to-dateness. We define satisfying here as the capability to match a professionally produced guidebook. We were also interested in developing and testing a method for empirical assessment of up-to-dateness in both media. As discussed above, both media feature characteristics that should contribute to up-to-dateness. In a summative analysis we will test which combination of characteristics leads to a better up-to-dateness. We will not be able to explain the effects of single characteristics on the overall result.

We assume that the long publication cycles of the guidebook versus the much shorter time-to-publish of the online community favour the community and will outweigh the more professionally organized process of the guidebook publisher. Hence, our hypothesis is that *Online tourism communities have a higher up-to-dateness than printed travelguides*. The next chapter presents an empirical analysis of this hypothesis.

4. Methodology and Data collection

4.1. Data Set

To systematically compare the timeliness of printed guidebooks with the timeliness of online communities we assessed the up-to-dateness of information in each medium. As guidebook we chose a “Lonely Planet” guide which is widely used by independent travellers. For the online community we chose World66² which is a wiki community containing 79,923 articles on 21,862 travel destinations (according to World66).

As region for our evaluation we chose one of the most popular travel countries, namely Italy. We used the “Lonely Planet Italy” in its 6th edition (2004) which was the latest edition at the time of our evaluation. Correspondingly, we only used the section for Italy in World66. Data collection and verification took place between May 5th and June 8th 2005³.

For the evaluation of up-to-dateness we followed a two-step process. In a first step we collected an exploratory sample to gain insights into the character and the number of up-to-dateness relevant attributes of information items. This data sample consisted of a total of 328 objects with 164 objects chosen from the Lonely Planet Guide and 164 objects chosen from World66.

To gain a sample of objects we conducted interviews with travelers to generate a taxonomy with categories relevant for travelers. These categories included 1) General Information, 2) Administrative Information (e.g., Visa information), 3) Medical Information, 4) Accommodation and Gastronomy 5) Public Transport 6) Sightseeing and Entertainment 7) Leisure time, sports and shopping and 8) Education. These categories were used as reference to identify travel-relevant information items from the printed guidebook as well as from the wiki. Within each of these categories objects were drawn in a randomized manner from the two media, but we assured that we covered big cities and touristic locations throughout Italy.

Each object was linked to a number of different attributes which were assessable with reference to up-to-dateness (e.g. opening hours, e-Mail addresses, fax numbers, phone numbers etc.). The analysis yielded a total of 10 different attributes such as these.

Since the number of objects as well as the number of attributes were distributed unevenly across our categories we decided to draw a more controlled sub-sample in a second step. For this sample, we selected the three attributes for which a sufficiently big data set was available to assure a fair evaluation of both media and which allowed us to use sound statistical methods. This sample was confined to the attributes “address”, “phone number” and “price”.

Since the up-to-dateness of different attributes of the same object could be related to another (e.g. a restaurant moving to a new address may also change its phone number), we only assessed one attribute per object to ensure independent data values. This resulted in 40 objects for each of the three attributes (address, telephone number and price) as well as for each medium (printed guidebook and online community) resulting in a total of 240 evaluated objects. The decision which items of the original data set were to be selected for this more controlled data set was made by a random number generator to avoid a bias.

To rule out that our sub-sample favoured either the printed guidebook or the wiki in any way we calculated statistical inference tests for both our original sample as well as for our sub-sample.

4.2. Method

To evaluate the up-to-dateness we selected an object from either the Lonely Planet Guide or from the online community and assessed whether this information was still up-to-date (e.g. the phone number of the Peggy Guggenheim Museum in Venice).

² <http://www.world66.com/>

³ The wiki might be subjected to slight changes during the period of evaluation. However, considering our hypothesis this has no systematic effect on the results of the evaluation.

Information was verified by communication with the institution in question using phone calls or e-mail correspondence, or by checking on official websites. Some information was checked personally on the respective location and local travelguides were asked to check some of the elements. After a comparison of the information in the medium and the researched information, each object was categorized as either up-to-date or not up-to-date.

5. Results

While the categories served as a tool to identify travel-relevant information items the number of retrievable items for the different categories as well as for the two media differed decisively. Thus, a sound comparison of up-to-dateness rates across these categories were hardly possible. The comparison of timeliness across categories was not in the scope of this paper and requires further research. Table 2 shows the up-to-dateness rates for the two media across the eight categories as well as the number of retrieved attributes.

Table 2. Up-to-dateness rates for the two media across categories (in percent)

Category	Percentage Up-to-date Lonely Planet	Percentage Up-to-date World66
General Inf.	43	27
<i>N</i>	7	11
Administr. Inf.	79	81
<i>N</i>	24	16
Med. Inf.	100	80
<i>N</i>	11	5
Acomm. Gastron.	72	73
<i>N</i>	97	127
Public Transp.	60	84
<i>N</i>	42	19
Sights. Entertainm.	61	66
<i>N</i>	114	109
Leisure Time, Sport, Shopping	71	63
<i>N</i>	42	19
Education	50	94
<i>N</i>	12	16
TOTAL	66	70
<i>N</i>	349	322

Since the number of retrievable objects as well as the number of total attributes varied across categories we decided to run a more controlled study with a sub-

sample containing only three attributes with sufficient frequency for a statistical comparison.

Figure 2 shows the percentage of items which were up-to-date for the three attributes “address”, “phone number” and “price” as well as for the two media “printed guidebook” and “wiki community”. For the Lonely Planet guidebook the percentage of correct addresses amounts to .92, the percentage of correct phone number to .75 and the percentage of correct prices amounts to .22. For the wiki community the percentage of correct addresses amounts to .90, the percentage of correct phone numbers amounts to .80 and the percentage of correct prices amounts to .25. A clear difference can be seen between the two attributes “address” and “phone number” and the attribute “price”. We assume that this relates to the fact that prices usually change much more often than addresses and phone numbers, resulting in lower up-to-dateness values.

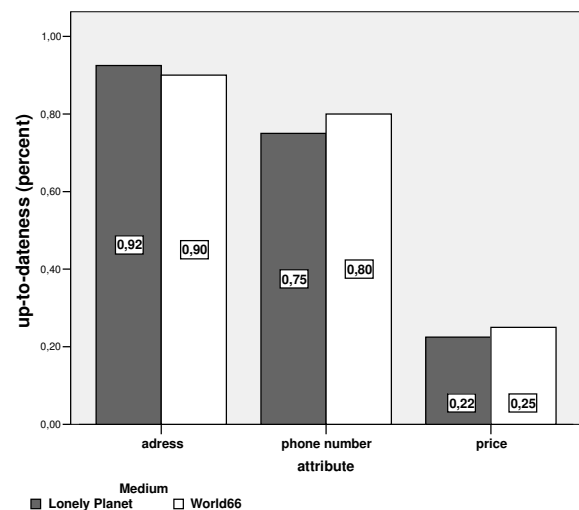


Figure 2. Proportional up-to-dateness for the three attributes as well as for the two media

For the two media, however, the frequencies are very similar. There is a slight descriptive advantage of the online community for the attributes “phone number” and “price” and a slight advantage of the printed guidebook for the attribute “address”.

In accordance with [11] we conducted a 3x2 analysis of variance with the two factors “attribute” (address, phone number, price) and “medium” (Lonely Planet, World66). The factor “medium” did not turn out to be significant ($p=.738$) whereas the factor “attribute” turned out to be highly significant ($p<.001$). The interaction also failed to reach significance ($p=.822$). Table 3 shows the result of this analysis.

To rule out any possibility that the main results were distorted by the fact that we did not calculate up-to-dateness values over our entire original data set, we also did an overall analysis of the original set which did not have an even number of attributes or objects for the two media. 66.76% of the objects from Lonely Planet were up-to-date (N=349) compared to 70.81% of the objects from World66 (N=322). An independent sample t-test did not result to be significant with $t(669) = -1.128$. Thus, we can conclude that a systematic difference between the media cannot be found for the controlled sub-sample nor for our original data set.

Table 3. Result of the analysis of variance

Source of Variation	SS	df	MS	F	Sig.
Model	119.25	6	19.20	133.8	.000
Attribute	20.36	2	10.12	68.55	.000
Medium	.017	1	.017	.112	.738
Atr*Med	.058	2	.029	.196	.822
Error	34.75	234	.149		
Total	154	240			

Hence, we have to reject our hypothesis that the online tourism community has a higher up-to-dateness than the printed travelguide. However, our data strongly indicates that the online communities are at least on the same level of up-to-dateness. To gain insight on the validity of this assumption we calculated the test power of our ANOVA factor "Medium". The power amounts to λ ($df = 1$, $\alpha = .05$) = 26, 67 for an effect of $\Omega^2 = .1$. This results in the probability of committing a Type II error of $\beta < .01$. Hence, the probability for the online community having a lower up-to-dateness compared to the guidebook is exceedingly low.

6. Interpretation

Our analysis shows that the printed guidebook and the online community can be claimed to have the same level of up-to-dateness. This result holds for the overall calculation of the up-to-dateness values as well as for each of the three attributes "address", "phone number" and "price". Whereas addresses and phone numbers can be considered reasonably up-to-date (75% or more) the values are dropping considerably for the attribute "price" (25% or less) in both media. This aspect is of interest because it might prove fruitful to systematically investigate attributes of different volatility. Even though our analysis did not yield this

result an evaluation with more attributes might show interesting differences. Online communities might be more suited for fast changing attributes or coordination activities of travelers, whereas for more stable attributes the printed guidebook could be as good as an online community.

Thus, the observed non-profit self-organizing online community can keep up in terms of up-to-dateness with a centralized professional content provider like a commercial travel guidebook. This similarity could stem from growing similarity in the information creation process: guidebook publishers increasingly tend to gather information from their clients, collecting feedback via email or hosting online communities with large discussion boards (cf. Lonely Planet's Thorn Tree). As we analyzed just one instance of each medium for one country, further research is needed for generalization of results and to gain an understanding of timeliness. We are currently conducting a study of up-to-dateness in several online communities and guidebooks for different countries.

7. Conclusion

We have presented a design for empirically assessing timeliness from a user perspective in online communities and information products and applied it to compare a printed guidebook and an online community in terms of up-to-dateness. In this setting we found that information from the online tourism communities was as up-to-date as the printed guidebook.

While previous approaches to assessing up-to-dateness have often concentrated on measures which could be assessed within an information system, like the data in a master compared to the data in a slave or the data in a cache compared to the data in the source (cf. [2]), we evaluated up-to-dateness by comparing the information in a system (or book) with the real world source from which the data was extracted.

The second interesting conclusion from this study is that information in a commercial professional information product is more outdated than one may expect. This might suggest that a habitual trust in organizational and institutional structures may not be justified.

This study leads to many questions for further research: What are the factors inhibiting higher up-to-dateness in online communities? A deeper understanding of the influence of community activity may help to determine if there is an inherent limit for up-to-dateness, possibly due to technological deficiencies or social processes. This may lead to

improvements for community support tools. As we have only analysed one online community, our findings need to be confirmed for other online communities. We strongly believe that the low Time-to-publish and the spread of mobile internet devices will lead to other improvements in social information systems and new usage patterns for online communities, e.g. regarding ad-hoc information needs and coordination tasks.

We see a need for further assessments of information quality from a user perspective, as many existing frameworks are based on a production perspective. As few studies on timeliness of information products and services have been published, there is still little understanding of what aspects of timeliness are relevant and how timeliness may be improved. We also observe a lack of studies on customers' perception of outdated information in terms of how many problems are caused by lower information quality.

Acknowledgements

We are grateful to Dustin Kerner. This research was partly funded by the Swiss National Science Foundation.

References

- [1] Alter, S., *Information Systems - A Management Perspective*, Addison-Wesley, Reading, MA, 1992.
- [2] M. Bouzeghoub and V. Peralta, "A framework for analysis of data freshness", In *Proceedings of the International Workshop on Information Quality in Information Systems (IQIS'04)*, pp. 59–67, 2004.
- [3] M. Burgess, N. J. Fiddian, and W. Gray, "Quality measures and the information consumer", In *Proceedings of the Ninth International Conference on Information Quality (IQ 2004)*, 2004.
- [4] C. Cappiello, C. Francalanci, and B. Pernici, "Data quality assesment from the user's perspective", In *Proceedings of the International Workshop on Information Quality in Information Systems (IQIS'04)*, pp. 68–73, 2004.
- [5] Eppler, M., *Managing Information Quality: Increasing the Value of Information in Knowledge-intensive Products and Processes*, Springer, Heidelberg, 2003.
- [6] Frey, B.S., *Not Just for the Money: An Economic Theory of Personal Motivation*, Edward Elgar, Cheltenham, UK and Brookfield, USA, 1997.
- [7] Z. J. Gackowski, "Logical interdependence of data/information quality dimensions - A purpose-focused view on IQ", In Chengalur-Smith, S., Raschid, L., Long, J., and Seko, C., (Eds.), *Proceedings of the Ninth International Conference on Information Quality (ICIQ 2004)*, MIT, Cambridge, MA, 2004.
- [8] Hagel, J. and Armstrong, A., *Net Gain: Expanding Markets Through Virtual Communities*, Harvard Business School Press, Harvard, 1997.
- [9] B.K. Kahn, D.M. Strong, and R.Y. Wang, "Information quality benchmarks: Product and service performance", *Commun. ACM* 45(4), pp. 184-192, 2002.
- [10] Kuhlen, R., *Mondlandung des Internet. Elektronische Kommunikationsforen im Bundestagswahlkampf 98*, Universitätsverlag Konstanz (UVK), Konstanz, 1998.
- [11] G.H. Lunney, "Using analysis of variance with a dichotomous dependent variable: An empirical study", *Journal of Educational Measurement*, 7 (4), 1970.
- [12] A. Neus, "Managing Information Quality in Virtual Communities of Practice", In Pierce, E. & Katz-Haas, R. (Eds.) *Proceedings of the 6th International Conference on Information Quality at MIT*. Boston, MA: Sloan School of Management, 2001.
- [13] M. Osterloh, B.S. Frey, and J. Frost, "The Dynamics of Motivation on New Organizational Forms", *International Journal of the Economics of Business, Special Issue on New Organizational Forms*, Vol. 9/1, 2002, pp. 61-77.
- [14] Preece, J. *Online Communities: Designing Usability and Supporting Sociability*, John Wiley & Sons, Chichester, 2000.
- [15] M. Prestipino, "Supporting Collaborative Information Spaces for Tourists", In R. Keil-Slawik, H. Selke, G. Szwillus, (Eds.), *Proceedings Mensch & Computer 2004*, Paderborn, 5. – 8. September 2004.
- [16] M. Prestipino, F.-R. Aschoff, and G. Schwabe, "What's the use of guidebooks in the age of collaborative media? Empirical evaluation of free and commercial travel information", In *Proceedings of the 19th Bled eConference eValues*. Bled, Slovenia, 2006.
- [17] Rheingold, H., *The Virtual Community*. Perseus Books, Jackson, 1993.
- [18] M., Rittberger, "Quality evaluation of electronic communication fora with evalYOUate", In Klein, B. and Rossin, D., (Eds.), *Proceedings of the Conference on Information Quality*, M.I.T., 2000, pp. 137–147.
- [19] M. Scannapieco and C. Batini, "Completeness in the relational model: A comprehensive framework", In

Proceedings of the Ninth International Conference on Information Quality (ICIQ-04), 2004, pp. 333-345.

[20] G. Schwabe and M. Prestipino, "How Tourism Communities can change travel information quality", In *Proceedings of the 13th European Conference on Information Systems (ECIS)*, Regensburg, Germany, 2005.

[21] Senn, J. A., *Information Systems in Management*. Wadsworth, Belmont, CA, 4th edition, 1990.

[22] R. Stockdale, M. Borovicka, "Developing an online business community: A travel industry case study", In *Proceedings of the 39th Hawaii International Conference on System Sciences (HICCS)*, Hawaii, 2006.

[23] Wenger, E., McDermott, R. and Synder, W., *Cultivating Communities of Practice*, Harvard Business School Press, Boston, 2002.